

**IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE**

NUANCE COMMUNICATIONS, INC., :
: Plaintiff, :
: v. : : Civ. No. 09-585-LPS
: :
VLINGO, CORP. : :
: Defendants, :
:

MEMORANDUM OPINION

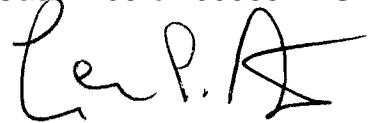
Jack B. Blumenfeld, Rodger D. Smith II, MORRIS, NICHOLS, ARSHT & TUNNELL LLP, Wilmington, DE; Morgan Chu, Jonathan Kagan, IRELL & MANELLA LLP, Los Angeles, CA; David McPhie, Rebecca Clifford, IRELL & MANELLA LLP, Newport Beach, CA,

Attorneys for Plaintiff.

Melanie K. Sharp, Erika R. Caesar, YOUNG CONAWAY STARGATT & TAYLOR, LLP, Wilmington, DE; Paul J. Hayes, Dean G. Bostock, MINTZ, LEVIN, COHN, FERRIS, GLOVSKY AND POPEO P.C., Boston, MA,

Attorneys for Defendants.

September 7, 2011
Wilmington, DE



STARK, U.S. District Judge:

Plaintiff Nuance Communications, Inc. (“Nuance”) filed the instant lawsuit on August 7, 2009, alleging that Defendant Vlingo Corporation (“Vlingo”) infringes three Nuance patents. (D.I. 1) On October 30, 2009, Vlingo answered and counterclaimed, asserting that Nuance infringes one of Vlingo’s patents. (D.I. 11) Presently pending before the Court is the matter of claim construction. The Court held a claim construction hearing on June 20, 2011. (D.I. 97, hereinafter “Tr.”) This is the Court’s Opinion on the proper construction of the terms in dispute.

I. INTRODUCTION

There are four patents involved in this lawsuit (collectively, the “patents-in-suit”). Nuance owns three of the patents: U.S. Patent No. 5,799,279 (“’279 patent”); U.S. Patent No. 6,871,179 (“’179 patent”); and U.S. Patent No. 7,315,818 (“’818 patent”). Vlingo owns the rights in the fourth patent: U.S. Patent No. 7,069,213 (“’213 patent”). Each of the patents-in-suit relates generally to the field of art involving speech recognition. Speech recognition is the process by which words spoken by a user are converted into a set of written words (or transcribed) by a computer or other device. (’179 patent, col. 1 lines 21-22) A speech recognition system receives a user’s dictated words in the form of speech signals; the speech signals are processed using complicated algorithms and “recognized” by identifying a corresponding text phrase in a vocabulary database. The recognized speech is then converted into an application where it can be displayed as text and used in many ways in a variety of software applications. (*Id.* at col. 1 lines 32-40) For example, speech recognition systems can recognize words spoken by a user and insert the dictated text into a letter or other document. (*Id.* at col. 1 line 29-30) While the patents-in-suit have differences, each is fundamentally based on

these general concepts.

II. LEGAL STANDARDS

“It is a bedrock principle of patent law that the claims of a patent define the invention to which the patentee is entitled the right to exclude.” *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312 (Fed. Cir. 2005) (internal quotation marks omitted). Construing the claims of a patent presents a question of law. *See Markman v. Westview Instruments, Inc.*, 52 F.3d 967, 977-78 (Fed. Cir. 1995), *aff’d*, 517 U.S. 370, 388-90 (1996). “[T]here is no magic formula or catechism for conducting claim construction.” *Phillips*, 415 F.3d at 1324. Instead, the court is free to attach the appropriate weight to appropriate sources “in light of the statutes and policies that inform patent law.” *Id.*

“[T]he words of a claim are generally given their ordinary and customary meaning . . . [which is] the meaning that the term would have to a person of ordinary skill in the art in question at the time of the invention, i.e., as of the effective filing date of the patent application.” *Id.* at 1312-13 (internal citations and quotation marks omitted). “[T]he ordinary meaning of a claim term is its meaning to the ordinary artisan after reading the entire patent.” *Id.* at 1321 (internal quotation marks omitted). The patent specification “is always highly relevant to the claim construction analysis. Usually, it is dispositive; it is the single best guide to the meaning of a disputed term.” *Vitronics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576, 1582 (Fed. Cir. 1996).

While “the claims themselves provide substantial guidance as to the meaning of particular claim terms,” the context of the surrounding words of the claim also must be considered. *Phillips*, 415 F.3d at 1314. Furthermore, “[o]ther claims of the patent in question, both asserted and unasserted, can also be valuable sources of enlightenment . . . [b]ecause claim terms are

normally used consistently throughout the patent.” *Id.* (internal citation omitted).

It is likewise true that “[d]ifferences among claims can also be a useful guide For example, the presence of a dependent claim that adds a particular limitation gives rise to a presumption that the limitation in question is not present in the independent claim.” *Id.* at 1314-15 (internal citation omitted). This “presumption is especially strong when the limitation in dispute is the only meaningful difference between an independent and dependent claim, and one party is urging that the limitation in the dependent claim should be read into the independent claim.” *SunRace Roots Enter. Co., Ltd. v. SRAM Corp.*, 336 F.3d 1298, 1303 (Fed. Cir. 2003).

It is also possible that “the specification may reveal a special definition given to a claim term by the patentee that differs from the meaning it would otherwise possess. In such cases, the inventor’s lexicography governs.” *Phillips*, 415 F.3d at 1316. It bears emphasis that “[e]ven when the specification describes only a single embodiment, the claims of the patent will not be read restrictively unless the patentee has demonstrated a clear intention to limit the claim scope using words or expressions of manifest exclusion or restriction.” *Liebel-Flarsheim Co. v. Medrad, Inc.*, 358 F.3d 898, 906 (Fed. Cir. 2004) (internal quotation marks omitted), *aff’d*, 481 F.3d 1371 (Fed. Cir. 2007).

In addition to the specification, a court “should also consider the patent’s prosecution history, if it is in evidence.” *Markman*, 52 F.3d at 980. The prosecution history, which is “intrinsic evidence,” “consists of the complete record of the proceedings before the PTO [Patent and Trademark Office] and includes the prior art cited during the examination of the patent.” *Phillips*, 415 F.3d at 1317. “[T]he prosecution history can often inform the meaning of the claim language by demonstrating how the inventor understood the invention and whether the inventor

limited the invention in the course of prosecution, making the claim scope narrower than it would otherwise be.” *Id.*

A court also may rely on “extrinsic evidence,” which “consists of all evidence external to the patent and prosecution history, including expert and inventor testimony, dictionaries, and learned treatises.” *Markman*, 52 F.3d at 980. For instance, technical dictionaries can assist the court in determining the meaning of a term to those of skill in the relevant art because such dictionaries “endeavor to collect the accepted meanings of terms used in various fields of science and technology.” *Phillips*, 415 F.3d at 1318. In addition, expert testimony can be useful “to ensure that the court’s understanding of the technical aspects of the patent is consistent with that of a person of ordinary skill in the art, or to establish that a particular term in the patent or the prior art has a particular meaning in the pertinent field.” *Id.* Nonetheless, courts must not lose sight of the fact that “expert reports and testimony [are] generated at the time of and for the purpose of litigation and thus can suffer from bias that is not present in intrinsic evidence.” *Id.* Overall, while extrinsic evidence “may be useful” to the court, it is “less reliable” than intrinsic evidence, and its consideration “is unlikely to result in a reliable interpretation of patent claim scope unless considered in the context of the intrinsic evidence.” *Id.* at 1318-19.

Finally, “[t]he construction that stays true to the claim language and most naturally aligns with the patent’s description of the invention will be, in the end, the correct construction.” *Renishaw PLC v. Marposs Societa’ per Azioni*, 158 F.3d 1243, 1250 (Fed. Cir. 1998). It follows that “a claim interpretation that would exclude the inventor’s device is rarely the correct interpretation.” *Osram GmbH v. Int’l Trade Comm’n*, 505 F.3d 1351, 1358 (Fed. Cir. 2007). Thus, if possible, claims should be construed to uphold validity. *See In re Yamamoto*, 740 F.2d

1569, 1571 (Fed. Cir. 1984).

III. CONSTRUCTION OF DISPUTED TERMS

The parties dispute a total of fourteen terms in the four patents-in-suit. The Court addresses each, in turn, below.¹

A. '279 Patent²

1. “continuous speech”

Claims 1 and 22 of the '279 patent recite, “A method for use in recognizing *continuous speech* . . .” ('279 patent, col. 11 lines 17-18 (emphasis added)) Nuance contends that the term should be construed as “speech not requiring the user to switch between separate modes.” (D.I. 71 at 4) Vlingo’s proposed construction for the term is “a series of words spoken by a user.” (*Id.*)

The Court has an obligation to resolve the parties’ dispute as to the meaning of this claim term. *See O2 Micro Intern. Ltd. v. Beyond Innovation Tech. Co., Ltd.*, 521 F.3d 1351, 1362 (Fed. Cir. 2008) (“A determination that a claim term ‘needs no construction’ or has the ‘plain and ordinary meaning’ may be inadequate when a term has more than one ‘ordinary’ meaning or when reliance on a term’s ‘ordinary’ meaning does not resolve the parties’ dispute.”). *O2 Micro* explained that while district courts are not required to engage in an exercise of redundancy, nor

¹Nuance contends that nearly every disputed claim term in its three patents ('279 patent, '179 patent, and '818 patent) does not actually require construction. For most of these terms, Nuance submits alternative constructions, in the event that the Court decides the term does require construction. Unless stated otherwise, the Court treats the alternative proposals provided by Nuance as Nuance’s proposed construction.

²The '279 patent is entitled “Continuous speech recognition of text and commands.” (D.I. 1 Ex. B)

to construe every single limitation in a claim, when the parties do present a fundamental dispute regarding the scope of a claim term, “it is the court’s duty to resolve it.” *Id.* In this case, the claim dispute needs to be resolved.

The central dispute relates to whether the “continuous speech” term embodies the distinction between the command mode and the dictation mode. The ’279 patent explains that, in the prior art, “[m]any speech recognition systems recognize spoken text in one mode and spoken commands in another mode.” (’279 patent, col. 1 lines 6-7) The two modes can be categorized as dictation mode and command mode. Dictation mode operates with the user speaking words without having to follow any particular structure; the system then transcribes words that are in its vocabulary. (*Id.* at col. 1 lines 10-14) The command mode, on the other hand, typically uses structured speech to instruct the system to take some action. For example, in the command mode, the user may say “underline last three words.” (*Id.* at col. 1 lines 20-24) Speech recognition systems often require the user to take some action – either clicking on an icon or verbally instructing the system – to switch between the two modes.

Nuance takes issue with Vlingo’s proposed construction for “continuous speech” insofar as it does not clarify that the patent recognizes spoken commands without requiring the user to take some action to switch between separate modes. (D.I. 71 at 5) The Court agrees with Nuance to the extent that the specification makes clear that the patented invention functions without requiring the user to switch between different modes. (’279 patent, col. 2 lines 6-14: “Recognizing spoken commands within dictated text allows users to intermittently execute commands that affect the text without requiring the user to switch between separate command and dictation modes.”) The Court does not agree with Nuance to the extent Nuance believes the

“without switching between modes” feature of the invention is at issue in the construction of “continuous speech.”³

Both parties’ proposed constructions are, therefore, flawed. Nuance’s is off because it is directed to an issue not implicated by the term “continuous speech.” Vlingo’s construction is problematic because it defines “speech” in a manner that would render “continuous” superfluous.

In order to construe this term, the Court finds that the ’952 patent is helpful and highly instructive. The ’279 patent incorporates by reference the ’952 patent and explains that the ’952 patent has a “more detailed description of continuous speech recognition.” (’279 patent, col. 3 lines 56-60) The ’952 patent explains that continuous speech is “the type of [speech] which humans normally speak,” in which “words are run together, without pauses or other simple cues to indicate where one word ends and the next begins.” (D.I. 59 Ex. E; ’952 patent, col. 3 lines 54-57) At oral argument, Nuance indicated that the ’952 description of continuous speech “is consistent with the [’279] patent, which is you don’t require the stilted form of communication, saying ‘Email.’ (Pause.) ‘Mother.’ (Pause.)” (Tr. at 11) Vlingo, too, agreed that the ’952 patent description of continuous speech is appropriate. (Tr. at 72: “Vlingo would be perfectly happy if the Court adopted that sentence from the ’952 patent”)

Accordingly, the Court will construe “continuous speech” to mean “the type of speech that humans normally speak, in which words are run together, without pauses or other simple cues to indicate where one word ends and the next begins.”

³It is unclear whether Nuance has this belief. At the hearing, Nuance’s counsel stated that how the invention is different than the prior art as it relates to switching between dictation and command modes is “not necessarily in claim 1.” (Tr. at 9)

2. “designating the recognized speech element as corresponding to a text element or to a command element”

The next disputed term appears in claims 1, 20, and 22 of the '279 patent. Nuance contends that this term does not require construction. (D.I. 71 at 5) Vlingo counters that the term should be construed as “assigning the characteristic of being a text or command element to a speech element, as opposed to merely recognizing that speech element is a text or command.” (D.I. 69 at 3)

The Court declines to import Vlingo’s requested “assigning” limitation into this term. Nor will the Court narrow the scope of this term so as to exclude situations in which the system knows – or is told – that the speech element is text or command, as opposed to the system figuring this information out itself. Vlingo’s position on these points is based on a purported prosecution history disclaimer. According to Vlingo, the patentee distinguished the Larkey reference by explaining that Larkey did not “designate” a speech element. Larkey merely recognized an utterance as a text or a command. Therefore, in Vlingo’s view, the “designating” limitation must mean something more than merely recognizing a speech element as text or a command. (D.I. 58 at 7)

Vlingo’s prosecution history argument fails, as Vlingo had not identified a clear, unmistakable disavowal of claim scope. *See Omega Eng’g, Inc. v. Raytek Corp.*, 334 F.3d 1314, 1324 (Fed. Cir. 2003) (explaining high standard for prosecution disclaimer argument); *see also McKesson Automation, Inc., v. Swisslog Italia S.P.A.*, 712 F.Supp.2d 283, 300 (D. Del. 2010). The portion of the prosecution history upon which Vlingo relies explains that the difference between Larkey and the patentee’s claim was that Larkey did not designate a speech element

based on the context in which the speech element arose, stating: “Larkey nowhere indicates that a particular recognition candidate [i.e., speech element] may be text in one context and a command in another context (e.g., the text ‘file’ and the command ‘file’).” (D.I. 59 Ex. F at 7-8) This is not a clear disavowal of a system that merely recognizes speech elements as text or speech. At best, any disclaimer here is ambiguous. *See generally, Rexnord Corp. v. Laitram Corp.*, 274 F.3d 1336, 1347 (Fed. Cir. 2001) (declining to limit term when prosecution disclaimer was inconclusive).

Having determined that there is no prosecution disclaimer relating to this disputed claim term, the Court agrees with Nuance that no further construction is necessary.

3. **“context”/“based on a context”**

Claim 1 recites, “wherein a particular one of the speech elements may correspond to a text element in one context and to a command element in another context . . .” (’279 patent, col. 11 lines 22-25) Nuance provides as its proposal for this term, “surrounding circumstances, which could include surrounding speech or text.” (D.I. 56 at 10) Vlingo contends that the term means “a particular structure, as defined by a template.” Vlingo further argues that “based on a context” means “adhering to a particular structure, as defined by a template.” (D.I. 69 at 5)

Vlingo argues initially that this term is indefinite under 35 U.S.C. § 112 and, therefore, incapable of being construed. (D.I. 58 at 8-9) At this stage of the proceedings, however, Vlingo falls far short of demonstrating by clear and convincing evidence that the claim term “context” is insolubly ambiguous, such that it is incapable of construction. *See generally Halliburton Energy Servs. v. M-I LLLC*, 514 F.3d 1244, 1249-1250 (Fed. Cir. 2008) (describing indefiniteness as “exacting standard,” which is satisfied when claim term is “completely dependent on a person’s

subjective opinion”); *Exxon Res. & Eng’g Co. v. United States*, 265 F.3d 1371, 1375 (Fed. Cir. 2001) (“If the meaning of the claim is discernible, even though the task may be formidable and the conclusion may be one over which reasonable persons will disagree, we have held the claim sufficiently clear to avoid invalidity on indefiniteness grounds.”).⁴

The Court adopts Nuance’s construction and, hence, will construe the term “context” to mean “surrounding circumstances, which could include surrounding speech or text.” This construction is consistent with the teaching of the specification. By contrast, Vlingo’s construction would result in a situation in which both command elements and text elements are defined by templates. The intrinsic record, however, provides support only for command templates, not also text templates. (’279 patent, col. 4 lines 45-65 (describing command templates)) Finally, Vlingo admits that under its construction, claim 6 and claim 8 would have the very same scope, running afoul of the doctrine of claim differentiation, as claim 8 depends from claim 6.⁵ (Tr. at 91)

Similarly, the Court sees no need to import an “adhering” limitation into the term “based on a context,” as Vlingo proposes. Therefore, the Court will construe “based on a context” as “based on surrounding circumstances, which could include surrounding speech or text.”

⁴On each of the disputed terms, the Court finds it is able to provide a construction, to the extent necessary. *See generally Pharmastem Therapeutics, Inc. v. Viacell, Inc.*, 2003 WL 124149, at *1 n.1 (D. Del. Jan. 13, 2003).

⁵Vlingo’s other objections to Nuance’s construction go to invalidity concerns, such as lack of antecedent basis for the third use of “context” in the claim (raised for the first time at the hearing) and lack of enablement/written description. (Tr. at 79-81, 88) These issues may all be raised at the summary judgment stage of this litigation. For present purposes, it is enough to say that the Court has found it possible to construe these disputed claim terms.

4. **“command element”**

The innovation in the '279 patent involves the speech recognition system's ability to distinguish when a user is directing a device to take an action – a command – and when a user is simply dictating text to be transcribed. Thus, claim 1 of the '279 patent teaches “accepting signals corresponding to interspersed speech elements including *text elements* corresponding to text to be recognized and *command elements* corresponding to commands to be executed.” ('279 patent, col. 11 lines 19-23 (emphasis added)) Vlingo proposes that “command element” be construed as “a word or phrase defined by a template that corresponds only to an action to be performed by a computer.” (D.I. 58 at 10; Tr. at 97-98) Nuance submits that the term should be construed as “a speech element corresponding to an order for an action to take place.”⁶ (D.I. 56 at 12)

The Court agrees with Nuance. Nuance's proposed construction is consistent with the claim language as well as the specification and does not improperly import any limitation from a preferred embodiment into the claim. Vlingo's construction, by contrast, would require a computer to function, but nothing in the claims or the specification requires that the invention be performed on a computer. Likewise, Vlingo's construction requires that a “command element” be “defined by a template,” which would limit the claims to a preferred embodiment. (See '279 patent, col. 4 line 47) (explaining that “in one method . . .”)

⁶Nuance also proposes that the Court could simply construe the more general term, “command,” as that term is common to both “command element” and “command template.” (D.I. 56 at 12) The Court disagrees. Construing “command” would not necessarily resolve the dispute for the terms “command element” and “command template.”

5. **“command template”**

Claim 8 of the '279 patent recites, “determining if the recognized speech element conforms to a command template.” ('279 patent, col. 11 lines 55-56) Vlingo contends that “command template” should be construed as “a list of command words and a corresponding order of the same that can be used to identify command elements.” (D.I. 58 at 11) Nuance contends that the term means “a rule or pattern relating to an available command.” (D.I. 56 at 12) The Court agrees with Nuance and will construe command template as “a rule or pattern relating to an available command.”

Both parties cite to the same section of the specification and argue that it supports their construction. (D.I. 69 at 10) The section provides:

The CPU then determines (steps **62** and **64**) whether the results represent a command or text. Commands include single words and phrases that are defined by templates (i.e., restriction rules). The templates define the words that may be said within command sentences and the order in which the words are spoken.

('279 patent, col. 4 lines 49-55) Vlingo's construction for “command template” improperly imports limitations from a preferred embodiment. The claim language itself does not require that a command template must be a list of command words and a corresponding order. Instead, the specification explains that a template is a “restriction rule,” and that the CPU compares the recognition results to possible command words and phrases and to command templates to determine if the recognition results are text or commands. ('279 patent, col. 4 lines 50-65) Nor does the Court agree with Vlingo that the patentee acted as its own lexicographer with respect to this term. (See '279 patent, col. 4 lines 52-53; Tr. at 92-93)

6. “acting on the recognized speech”

Claim 1 provides, “acting on the recognized speech elements in a manner which depends on whether the speech elements correspond to text elements or command elements.” (’279 patent, col. 11 lines 34-36) Vlingo contends that “acting on the recognized speech” should be construed as “providing text to a text processing application or causing an application to cause a step.” (D.I. 58 at 11) Nuance, on the other hand, contends that this term does not require construction and does not provide an alternative. (D.I. 56 at 14)

This dispute relates to dependent claims 2 and 3. Claim 2 recites, “The method of claim 1 in which a text element is acted on by providing the text element to a text processing application.” (’279 patent, col. 11 lines 37-39) Claim 3 next recites, “The method of claim 1 in which a command element [sic] acted upon by causing an application to perform a step.” (*Id.* at col. 11 lines 40-41) Thus, as Nuance points out, Vlingo’s proposal for “acting on the recognized speech” repeats verbatim language contained in dependent claims 2 and 3.

Nuance contends that Vlingo’s additional language is an attempt to import limitations from the specification into the claims. The Federal Circuit has repeatedly warned against importing limitations from a preferred embodiment into the claims of a patent-in-suit. *See, e.g., Phillips*, 415 F.3d at 1323 (“For instance, although the specification often describes very specific embodiments of the invention, we have repeatedly warned against confining the claims to those embodiments.”); *Liebel-Flarsheim Co.*, 358 F.3d at 908 (“Even when the specification describes only a single embodiment, the claims of the patent will not be read restrictively unless the patentee has demonstrated a clear intention to limit the claim scope using ‘words or expressions of manifest exclusion or restriction.’” (internal citations omitted)). Here, there is no “clear

intention to limit the claim scope using words or expressions of manifest exclusion or restriction.” *Liebel-Flarsheim Co.*, 358 F.3d at 906. Accordingly, the Court rejects Vlingo’s proposed construction.

At the same time, Nuance does not provide a construction for the term. In this case, however, some construction is needed. Figure 4 of the patent indicates the next step after “recognizing” the speech is to “compare results against command templates.” (’279 patent, Fig. 4) Comparing the recognized speech is acting on the recognized speech. Moreover, as Nuance argued at the hearing, there are other actions that a system could take with respect to recognized speech, such as buffering the speech element or storing the speech element as a wave file. (Tr. at 28) Vlingo’s construction is too narrow.

Accordingly, the the Court will construe “acting on the recognized speech” as “taking some action on the recognized speech, including providing text to a text processing application or causing an application to cause a step.”

7. **“recognizing . . . in parallel”**

Dependent claim 12 provides,

The method of claim 1 in which the recognizing comprises, in parallel

recognizing the speech elements as if they were text,
and

recognizing the speech elements as if they were
commands.

(’279 patent, col. 11 line 65 - col. 12 line 3) Nuance contends that this term should be construed as “recognizing . . . concurrently.” (D.I. 56 at 15) Vlingo contends that the correct construction

is “recognizing using more than one speech recognition engine at the same time.” (D.I. 58 at 12) The Court will construe this term as “recognizing . . . simultaneously.”

The central issue in dispute relating to this term is whether the claim may be practiced with a single recognition engine. (D.I. 56 at 15; D.I. 58 at 12) There is no support in the claim or the specification for requiring, as Vlingo’s construction does, that more than one speech recognition engine perform the recognition.⁷ The portion of the specification on which Vlingo relies actually undermines Vlingo’s contention. For example, in describing Figure 14, the ’279 patent states:

Another alternative is parallel speech recognition of both dictated text and commands. Referring to Fig. 14, the CPU simultaneously recognizes (steps 160 and 162) dictated text and commands by simultaneously comparing the speech frames of the user utterance against one or more dictated text vocabularies 150 and one or more command vocabularies 152.

(’279 patent, col. 9 lines 34-40) From this passage, Vlingo suggests that the specification does not provide support for a single speech recognition engine performing “simultaneous” speech recognition. The Court disagrees, finding here support for a single CPU simultaneously recognizing text and commands.

The specification’s use of “simultaneously” indicates that the recognition occurs at the same time. The Court does not agree with Vlingo that simultaneously requires two events start *and end* at the same time as one another. (Tr. at 100-01) Nothing in the claim or specification requires that the recognition processes must *end* at the same time. Instead, the specification

⁷Vlingo also presses the argument that, if the claim does not require more than one recognition engine, then it lacks sufficient written description or is not enabled under 35 U.S.C. § 112. (D.I. 58 at 12 n.5) The Court will not address invalidity contentions at this time.

expressly contemplates that one process may cease when it becomes apparent that the speech element is or is not a command. ('279 patent, col. 9 lines 52-58)

8. “re-recognition”

The '279 patent contemplates situations in which a user may make a correction if the speech recognition system incorrectly categorizes a speech element. Thus, claims 16, 17, and 22 provide that, under certain conditions, a user may “cause a re-recognition.” ('279 patent, col. 12 lines 12-15) Nuance argues that the term “re-recognition” should be construed as “recognizing an utterance for a second or more time.” (D.I. 56 at 16) Vlingo argues that the term should be construed as “recognizing a previous utterance for a second or more time.” (D.I. 58 at 13)

The Court will construe this term as “recognizing an utterance for a second or more time.” As is evident from the parties’ proposed constructions, the only difference is whether the utterance must be a “previous” utterance – that is, whether the user speaks the same utterance for a second time. (D.I. 69 at 12) Vlingo’s “previous” limitation, however, does not appear in the claims and would improperly limit the scope of the claim. Vlingo characterizes Nuance’s proposed construction as an attempt to argue that if the user speaks the same utterance for a second time, this second utterance is not a re-recognition, but a new recognition. (D.I. 69 at 12) The Court disagrees. Nuance’s construction does not require such a result.

B. '179 Patent⁸

1. “voice command”

The term “voice command” appears in independent claims 1 and 6 of the '179 patent,

⁸The '179 patent is entitled “Method and apparatus for executing voice commands having dictation as a parameter.” (D.I. 1 Ex. C)

both of which recite a “voice command having a voice command component.” (’179 patent, col. 8 line 44; *id.* at col. 9 line 13) Nuance proposes that this term should be construed as “a spoken order for an action to take place.” (D.I. 56 at 17) Vlingo contends that the proper construction of the term is “a spoken utterance that causes the system to perform a pre-determined function within system or application software other than transcribing text.”⁹ (D.I. 58 at 15) The Court will construe this term as “a spoken order for an action to take place.”

Vlingo’s construction is derived from the specification section entitled, “DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT.” (’179 patent, col. 4 lines 23-24; *id.* at col. 6 lines 22-25) Adopting it would have the effect of improperly importing a limitation of the preferred embodiment into the claims.¹⁰ *See Phillips*, 415 F.3d at 1320. Moreover, the claim language itself specifies that the “voice command” has both a “voice command component” and a “dictation component . . .” (’179 patent, col. 8 lines 44-45) Vlingo’s construction would appear to result in a voice command that cannot transcribe text, which is contradicted by the specification repeatedly. (*Id.* at col. 3 lines 4-6; *id.* at col. 3 lines 25-30; *id.* at col. 6 lines 58-60) According to the specification, “[a]lthough the system of the present invention can be used to recognize all three types of spoken utterances” – which the specification explains are “voice commands,” “ordinary dictation,” and “voice commands incorporating dictation” – “the invention is intended to address the unique difficulties in recognizing voice commands

⁹Vlingo initially argued that this term is indefinite under 35 U.S.C. § 112. (D.I. 58 at 15) At the hearing, Vlingo appeared to retreat from this position. (Tr. at 105-06)

¹⁰Vlingo also seems to argue that this term is another example of the patentee acting as its own lexicographer. (D.I. 69 at 15: “Nuance fails to . . . contradict[] this explicit definition of voice command.”) To the extent Vlingo makes such an argument, the Court disagrees with it.

incorporating dictation.” (’179 patent, col. 6 lines 31-34) Vlingo’s construction would also potentially introduce confusion into the claims, which undermines a central purpose of claim construction. *See Power-One v. Artesyn Techs., Inc.*, 599 F.3d 1343, 1348 (Fed. Cir. 2010) (explaining that construction must ensure that jury fully understands claims).

2. “command grammar”

Independent claims 1 and 6 recite “identifying a voice command having a voice command component . . . wherein said voice command component is specified by a command grammar . . .” (’179 patent, col. 8 lines 44-47) Nuance proposes that “command grammar” should be construed as “a set of available commands.” (D.I. 56 at 17) Vlingo contends that the Court should construe the term as “a finite set of structural rules or patterns that deterministically define commands.” (D.I. 58 at 16) The Court will construe command grammar as “a set of available commands,” which is consistent with the description of the invention in the specification and the prosecution history.

The specification explains that in a preferred embodiment, the system accepts spoken utterances by a user and compares the recognized words against a plurality of predetermined command patterns. The predetermined command patterns are contained in a command pattern “set[] or grammar[].” (’179 patent, col. 6 lines 49-54) The invention then matches the command grammar with a specified action. (*Id.* at col. 7 lines 31-35; *see also* D.I. 70 Ex. 3 at 4) The patent is explicit, however, that the above description is but one embodiment. (*Id.* at col. 7 lines 43-45)

The Court finds no support in the claims or specification for Vlingo’s “deterministically define” limitation. Vlingo’s construction attempts to narrow the scope of the claim by importing

limitations from the preferred embodiment, and the Court accordingly rejects it.

3. “free-form text”

Claim 1 and claim 6 recite a “dictation component . . . wherein . . . said dictation component is free-form text which is not specified by said command grammar.” (’179 patent, col. 8 lines 45-49; *id.* at col. 9 lines 14-18) Nuance contends that “free-form text” should be construed as “any text recognizable by the speech recognition system.” (D.I. 56 at 18) Vlingo asserts that the term should be construed as “text not in the command grammar.” (D.I. 58 at 17)

The Court will construe “free-form text” as “any text recognizable by the speech recognition system.” The claim itself provides substantial guidance, insofar as it specifies that the dictation component is free-form text that “is not specified by said command grammar.” (’179 patent, col. 6 lines 48-49) Moreover, the specification teaches that “[t]he dictation may be comprised of any set of words in a voice recognition vocabulary . . .” (’179 patent, col. 6 lines 58-68) Vlingo’s construction would result in a claim that read “dictation component is ***text not in the command grammar*** which is not specified by said command grammar.” Such a construction would be redundant.

4. “contiguous utterance”

Claims 1 and 6 recite a voice command having a voice command component and a dictation component “within a contiguous utterance.” (’179 patent, col. 8 lines 44-46; *id.* at col. 9 lines 14-15) Nuance contends that the term “contiguous utterance” should be construed as “a series of sounds spoken by a user.” (D.I. 56 at 19) Vlingo takes issue with “sounds,” arguing instead for a construction of “a series of words spoken by a user.” (D.I. 58 at 18)

The Court will construe contiguous utterance to mean “a series of words, including

acronyms and letters, spoken by a user.” This ruling captures aspects of both parties’ constructions, and addresses the concerns each party raised with the other’s proposal. Vlingo concedes that “utterances” can include more than just words, such as spoken numbers, letters, and acronyms. (D.I. 69 at 18) At the same time, Vlingo expressed concern that Nuance’s “sounds” would mean that the system could recognize “laughs, cries, sighs.” (Tr. at 112)¹¹

C. **'818 Patent**¹²

1. **“receiving an edited text”**

The '818 patent is directed toward correcting errors in speech recognition systems. Claim 1 of the '818 patent recites “receiving an edited text of the recognized speech . . .” ('818 patent, col. 31 lines 3-5) Nuance contends that “receiving an edited text” does not need to be construed. (D.I. 56 at 20) Vlingo, however, submits that the term is indefinite because it does not identify who edits the text. Alternatively, Vlingo argues that the Court should construe the term as “receiving an edited text from the user that dictated the original text.” (D.I. 58 at 13)

The Court declines to construe this term.¹³ Vlingo’s proposed construction tracks the language of the claim and then appends an additional limitation – namely, that the source of the edited text must be the same user that supplied the original text. (D.I. 58 at 14; D.I. 71 at 19) Vlingo’s construction, however, is an improper attempt to import a limitation from the specification. The Court does not agree with Vlingo that the specification leads to the

¹¹Vlingo seems to have abandoned a prosecution disclaimer argument.

¹²The '818 patent is entitled “Error correction in speech recognition.” (D.I. 1 Ex A)

¹³For the reasons previously noted, the Court does not, at this time, determine if the claim is invalid due to indefiniteness.

“inescapable conclusion” that the user submitting the edited text must be the same user that provided the original text. *See Microsoft Corp. v. Multi-Tech Sys.*, 357 F.3d 1340, 1347-48 (Fed. Cir. 2004) (applying principle that it is “improper to read a limitation from the specification” into claims).

In some respects, the specification teaches away from Vlingo’s construction. According to the specification, “the speech recognition system receives the user-dictated text by, for example, receiving a recording from a recorder or receiving user-dictated text directly through a microphone” (’818 patent, col. 26 lines 36-42) Here, the specification explicitly contemplates a situation in which a user may edit text from a previous recording; there is no indication that the person editing the text is the same user who spoke the text in the recording. Likewise, as the specification explains, “[i]n another implementation, the speech recognition system has already recognized the user-dictated text or speech and has stored it for later use in memory.” (*Id.* at col. 26 lines 36-38) Again, a different user may be involved in editing the recognized speech that the system has stored.

Vlingo effectively concedes that “receiving an edited text” has a plain and ordinary meaning, since its construction incorporates verbatim the claim language “receiving an edited text.” Having resolved the parties’ dispute as to whether to incorporate Vlingo’s additional limitation, the Court finds no need to further construe this term.

D. '213 patent¹⁴

1. “which one of the items from the list matches the utterance”

The '213 patent is directed toward “barge-in” speech recognition systems, which involve a user selecting from a list of items audibly provided to the user. The time the user “barges-in” to make an utterance influences the system’s recognition of the user’s speech. ('213 patent, col. 1 lines 51-61) Claim 1 of the '213 patent provides that the barge-in time “influence[s] the determination of which one of the items from a list matches the utterance.” ('213 patent, col. 9 lines 57-62) Nuance contends that the proper construction of the term in dispute is “which one of the items from the list has a sound that is most similar in comparison to the utterance.” (D.I. 56 at 6) Vlingo argues that the term has a plain meaning and therefore does not require any construction. (D.I. 58 at 19)

The Court agrees with Vlingo. The claim language and the specification do not support Nuance’s construction. For example, in a description of the preferred embodiment, the specification explains that the user may “say . . . the desired” item or, alternatively, “provide an utterance referring to” the desired item. ('213 patent, col. 4 lines 29-32) There would be no necessity of distinguishing “saying” the desired item from “providing an utterance referring to” the desired item unless the invention covered instances in which the utterance did not sound similar to the item on the list.

Nuance’s construction is based, effectively, on a single sentence from a description of the preferred embodiment that notes, “[t]hese commercially-available software products have the

¹⁴The '213 patent is entitled “Influencing a voice recognition matching operation with user barge-in time.” (D.I. 11 Ex. 1)

capability to compare an utterance with a given item and provide the degree of similarity of the two.” (’213 patent, col. 5 line 57 - col. 6 line 8) Nuance’s argument is flawed for two reasons: (1) this section is but one embodiment of the invention; and (2) the invention makes clear that it uses both the speech recognition software *and* the time the utterance was received, indicating that the invention will produce a confidence score, even if the utterance sounds nothing at all like any item on the list. (’213 patent, col. 1 lines 60-65) (explaining that confidence score is based on both similarity and barge-in time and that item is selected from list based on confidence score) Nuance improperly seeks to import into this claim term a limitation from a preferred embodiment.

Having resolved the parties’ sole dispute relating to this term, the Court finds it unnecessary to further construe it.

IV. CONCLUSION

For the foregoing reasons, the Court will construe the terms in the manner set forth in this Opinion. An appropriate Order will follow.